

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

REMARKS/ARGUMENTS

These remarks are submitted in response to the Final Office Action of April 28, 2005 (Office Action). This response is filed after the 3-month shortened statutory period, and as such, a retroactive extension of time is hereby requested. The Examiner is authorized to charge appropriate fees to Deposit Account 50-0951.

In paragraph 6 of the Office Action, Claims 1-4, 9-10, 12-13, 15-18, and 23 were rejected under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent No. 6,092,044 to Baker, *et al.* (hereinafter Baker). In paragraph 8 of the Office Action, Claims 5-6, 8, 14, 19-20, and 22 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Baker in view of U.S. Patent No. 6,363,342 to Shaw, *et al.* (hereinafter Shaw), and Claims 7, 11, and 21 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Baker in view of U.S. Patent No. 5,850,629 to Holm, *et al.* (hereinafter Holm).

Independent Claims 1, 10, 15 have been amended to further clarify certain features of Applicants' invention. Dependent Claims 2, 3, 5, 6, 16, 17, 19, and 20 have also been amended to emphasize certain aspects of the invention. The amendments, as described herein, are supported throughout the Specification. No new matter has been introduced by virtue of the amendments.

I. Applicants' Invention

Applicants' invention provides a computer-implemented method and a computer-based system for composing the pronunciation of a portion of text. According to one embodiment of the invention, a computer-implemented method of composing a pronunciation includes graphically presenting at least one activatable visual identifier corresponding to individual phonemes and generating pronunciation information in response to a user's selection of an activatable visual identifier. The pronunciation, more particularly, is generated in accordance with the activatable visual identifier selected and

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

comprises at least one phoneme selected from among multiple phonemes, an ordering of selected phonemes, a pronunciation stress parameter, and a prosodic parameter. (See, e.g., specification, p. 7, lines 8-14, and 24-26; p. 8, lines 13-20.) The pronunciation information can include, but is not limited to, features used by a speech recognizer to recognize speech. (Specification, p. 7, lines 7-10.)

The method, according to this embodiment, further includes enabling a user to compose the pronunciation of a portion of text by choosing to add a particular phoneme and/or remove a particular one. (Specification, p. 7, lines 15-21; p. 8, lines 11-24; p. 10, line 13- p. 11, line 3.) The user's choice for performing these operations, moreover, can be based upon the pronunciation information. The user's choice can also be based on an audible rendering of a portion of the pronunciation during the user's composing the pronunciation; that is, as the pronunciation develops and evolves, the user can initiate a playback of the pronunciation without compiling the pronunciation. (See, e.g., Specification, p.8, lines 1-3.)

The user's selection additionally, or alternatively, can be based on an audible rendering of an exemplary word illustrative of a particular phoneme and/or a visual rendering of an exemplary word illustrative of the particular phoneme. The method according to this embodiment of the invention also includes compiling pronunciation information responsive to a selection of one of the plurality of visual identifiers.

II. The Claims Define Over The Prior Art

As already noted, independent Claims 1, 10, and 15 were rejected as being anticipated by Baker. Baker is directed to a method of adding words to a speech recognition vocabulary. (Col. 1, lines 49-50; Abstract.) Baker's method of adding words to a speech recognition vocabulary begins with the "creation of a collection of possible phonetic pronunciations based on the spelling of [a] word," in which the collection "of

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

possible phonetic pronunciations is created by comparing the spelling of the word to a rules list of letter strings with associated phonemes." (Abstract; Col. 1, line 28 – Col. 2, line 12.) (Emphasis Supplied.) Speech recognition, including performing a conventional Fast Fourier Transform of digital samples of speech, is used to generate a pronunciation that is then used to "find a pronunciation from the collection that best matches a verbal utterance of the word." (Abstract; Col. 2, lines 13-21; See also Col. 3, line 65 – Col. 14, line 53; FIG. 2) (Emphasis Supplied.)

A. Baker's Comparison and Matching of Possible Pronunciations Is Not Equivalent to Composing a Pronunciation

Baker's building of a speech recognition vocabulary entails the generation of possible phonetic pronunciations. In this sense, Baker indeed generates a pronunciation. Beyond this semantic generality, however, there are fundamental differences between Baker and Applicants' invention. The differences are found in the way in which each generates a pronunciation.

As explicitly described throughout the reference, Baker generates a pronunciation by comparing word spellings with possible phonetic pronunciations and then selects from among the possible pronunciations the one that best matches a verbal utterance. The comparing and matching of already-processed pronunciations, however, is entirely distinct from "composing" a pronunciation phoneme-by-phoneme as with Applicants' invention. Baker precludes such composition since Baker relies on a pre-processed collection of pronunciations, none of which are individually constructed by a user's selectively adding and deleting individual phonemes, as explicitly recited in each of amended independent Claims 1, 10, and 15.

More particularly, the control window 1750 of Baker does not present activatable visual identifiers that correspond to individual ones of a plurality of phones that can be

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

used to compose a pronunciation by adding and deleting individual phonemes. What appears in Baker's pronunciation box 1756 of the control window 1750 is "the pronunciation of [a] word" that has already been "determined" by one of two methods. (Col. 18, lines 2-4.) The first method entails creating a constraint grammar containing a word list of possible phonetic spellings using a rules list. A recognizer 215 is then used to choose the best phonetic spelling based by comparing "the spoken word against the constraint grammar." (Col. 15, line 56 – Col. 17, line 31.) The presentation of an already-determined word pronunciation in Baker, obviates not only the need, but the opportunity, to compose the pronunciation from individual phonemes corresponding to activatable visual identifiers, as recited in amended independent Claims 1, 10, and 15.

Baker's second method of determining a word pronunciation that is presented in the pronunciation box 1756 entails using a rules list to create a "net" containing all possible phonetic spellings, the net being created after a user both spells a word and utters the word. (Col. 17, lines 31-36.) The recognizer 215 in Baker then uses continuous recognition of the uttered word to "prune out paths in the net" that do not contain phonemes corresponding to the spoken word. (Col. 17, lines 44-48.) Again, the result of Baker's second method is the presentation to a user of an already-determined word pronunciation. This second method, accordingly, also eliminates the opportunity for a user-composed pronunciation based on selective addition and/or removal of individual phonemes, as recited in amended independent Claims 1, 10, and 15.

B. Baker Provides Word Editing Not Phoneme Selection

Although Baker teaches editing a pronunciation presented in the pronunciation box 1756, Baker's editing does not entail each of the aspects of Applicants' invention. For example, editing in Baker does not encompass composing a pronunciation. In particular, Baker does not provide for composing a pronunciation by selectively adding

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

and/or removing individual phonemes, nor does Baker provide activatable visual identifiers corresponding to individual phonemes, as recited in each of amended independent Claims 1, 10, and 15.

That Baker does not encompass these aspects is apparent since their inclusion would render superfluous both of the methods employed by Baker to determine the pronunciation that is presented to a user in the pronunciation box 1756. More fundamentally, though, Baker's description of the control window 1750 in which the pronunciation box is displayed precludes these features.

As expressly described, a user activates a "word history" button 1770 to display words generated according to Baker. The activation allows the user to add and delete words themselves, not individual phonemes that correspond to activatable visual identifiers. No comparable addition or removal is even remotely suggested for individual phonemes, which, as already noted, are determined not by a user's composing a pronunciation but by comparing spelling-derived pronunciations of a word and finding the pronunciation that best matches an utterance of the word.

Moreover, because Baker precludes composing a pronunciation, phoneme-by-phoneme or otherwise, Baker does not permit composing a pronunciation at least partly based upon an audible rendering of a portion of the pronunciation during the user's composing the pronunciation without compiling pronunciation information. The text-to-speech button 1762 provided by the control window 1750 in Baker is limited to playing back phonemes in the pronunciation box 1756. As already noted, however, the phonemes presented by Baker in the pronunciation box 1756 are those that correspond to the pronunciation already determined by comparing possible pronunciations based on a spelling of a word and finding the one that best matches an utterance of the word. As there is no user-directed composing of the pronunciation with individual phonemes in

Appln. No. 10/007,615
Amendment dated Aug. 29, 2005
Reply to Final Office Action of April 28, 2005
Docket No. BOC9-2001-0039 (284)

Baker, there is neither a benefit from nor an opportunity for audibly rendering a portion of the pronunciation during the user's composing the pronunciation.

Applicants respectfully submit that Baker fails to disclose each of the features of amended independent Claims 1, 10, and 15, and that, therefore, the claims define over the prior art. Applicants further respectfully submit that whereas the remaining claims each depend from one of the amended independent claims while reciting additional features, these claims, too, define over the prior art.

CONCLUSION

Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: August 29, 2005



Gregory A. Nelson, Registration No. 30,577
Richard A. Hinson, Registration No. 47,652
AKERMAN SENTERFITT
Customer No. 40987
Post Office Box 3188
West Palm Beach, FL 33402-3188
Telephone: (561) 653-5000